Probabilidades y Estadística II
Modelización de la incertidumbre
Incertidumbre y Probabilidad

Indice

Modelización de la incertidumbre

Variables aleatorias

•El concepto de variable aleatoria permite pasar de los resultados experimentales (E) a una función numérica (real) de los resultados

•Dado el espacio probabílistico (E, $\mathcal{A}$, P), la aplicación ξ: E $\rightarrow \Re$, $\omega \rightarrow \xi(\omega)$ es un variable aleatoria si $\forall$ x $\in \Re$, $\{\omega \in$ E: $\xi(\omega) \leq x\}$ es un suceso, $\xi^{-1}((-\infty,x]) \in \mathcal{A}$, σ-Álgebra

•La probabilidad definida sobre sucesos se transforma en probabilidad de que la variable aleatoria ξ tome valores en $(-\infty,x]$, $P(\xi \leq x)$

•Se trata de un cambio de lenguaje: antes el algebra de Boole y ahora la Teoría de Funciones del Análisis (herramientas matemáticas: funciones de variable real, cálculo diferencial e integral,.....)

Modelización de la incertidumbre
Variables aleatorias

•Modelización: Abstracción

Modelo de distribución de probabilidad: especifiación de los posibles valores de la variable aleatoria con sus probabilidades

•Notación:

X, Y,... variables aleatorias. $X(\omega)=x$, $Y(\omega)=y$,....
número asociado al resultado $\omega \in E$, cuantificación

•*Lenguaje de sucesos $\rightarrow$ de probabilidad $\rightarrow$ de funciones reales*

Sucesos	Variable aleatoria
- Venta de un producto	número de productos vendidos
- Llegada de un cliente	número de clientes atendidos
- Tamaño de un e-mail	número de kbytes enviados por e-mail
- Fallo de un dispositivo	número de horas hasta el fallo de un dispositivo
- Curación de un paciente	número de años de supervivencia post-tratamiento
- Incendio forestal	número de hectáreas quemadas (+localización)

Modelización de la incertidumbre
Variables aleatorias

- Diferencias entre Estadística Descriptiva y Cálculo de Probabilidades:

  - La variable estadística es descriptiva, analiza hechos.

  - La variable aleatoria es probabilista, analiza causas potenciales, sobre el futuro, no hechos, el proceso generador de los hechos, datos

- Tipos de variables aleatorias: asociación entre resultado y un número real

  - Discreta: toma un conjunto de valores finito o infinito numerable
    Cardinal(E) $\rightarrow$ *N*

  - Continua: toma valores en un intervalo,
    Cardinal($\Re$) $\rightarrow$ potencia del continuo

Modelización de la incertidumbre

Variables aleatorias. Función de distribución

•La variable aleatoria no se presta al Análisis Real pues son funciones reales de sucesos (conjuntos) y no de variable real, sobre $\Re$

•Se introduce la función de distribución de la variable aleatoria ξ:

$$F: \Re \to [0,1],\ F(x) = P(\xi(\omega) \leq x)$$

•Propiedades:

1. $0 \leq F(x) \leq 1,\ \forall\ x \in \Re$
2. $\lim_{x \to -\infty} F(x) = 0,\ \lim_{x \to +\infty} F(x) = 1$
3. $x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2)$, monotonía no decreciente
4. $\lim_{x \to a+} F(x) = F(a),\ \forall\ a \in \Re$, continuidad por la derecha
5. $P(a \leq \xi \leq b) = F(b) – F(a)$, probabilidad de un intervalo

Modelización de la incertidumbre

Variables aleatorias discretas

- La variable aleatoria ξ se dice discreta si toma valores en un conjunto numerable $\{x_1,x_2,\ldots x_n,\ldots\}$, finito o infinito. Si $p_i = P(\xi \leq x_i) \geq 0$, i=1,2,…n,…

   1. $\Sigma_i \, p_i = 1$
   2. $P(\xi \leq x_n) = \Sigma^n{}_i \, p_i$

- Se define la función de probabilidad de la variable aleatoria ξ:

$$P(\xi=x) = P(\{\omega \in E: \xi(\omega)=x\})$$

- Asignación de probabilidad a los sucesos elementales sobre los que la variable aleatoria toma el valor x. Masa de probabilidad puntual

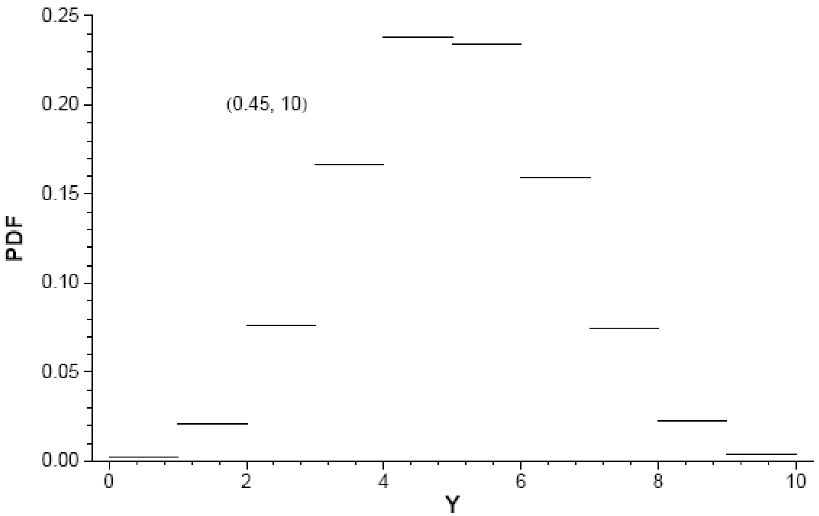
- Se obtiene la función de distribución de la variable aleatoria ξ:

$$F(x) = P(\xi \leq x_j) = P(\{\omega \in E: \xi(\omega) \leq x_j\}) = \Sigma_{xj \leq x} P(\xi=x_j) = \Sigma_{xj \leq x} P_j$$

La F(x) de una variable aleatoria discreta es escalonada

Modelización de la incertidumbre
Variables aleatorias discretas

Binomial(A,B) $y = 0, 1, 2, \ldots, \quad 0 < A < 1, \quad y, 2 \le B$



$$\text{PDF} = \binom{B}{y} A^{y} \left(1 - A\right)^{B-y}$$

Parameters -- A (p): Prob(success), B (N): Number of Bernoulli trials (constant)

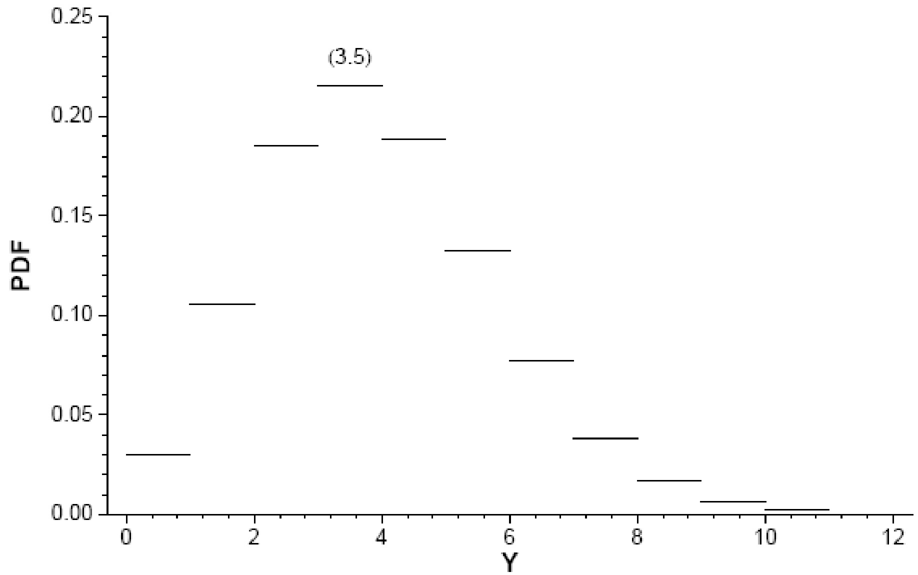
Moments, etc.

$$\text{Mean} = A\,B$$

$$\text{Variance} = A\left(1 - A\right)B$$

$$\text{Mode} = \text{int}\left(A\left(B + 1\right)\right)$$

Poisson(A) $y = 0, 1, 2, \ldots, \quad A > 0$



$$\text{PDF} = \frac{\exp\left(-A\right) A^{y}}{y!}$$

Parameters -- A (θ): Expectation

Moments, etc.

$$\text{Mean} = \text{Variance} = A$$

$$\text{Mode} = \text{int}\left(A\right)$$

Modelización de la incertidumbre

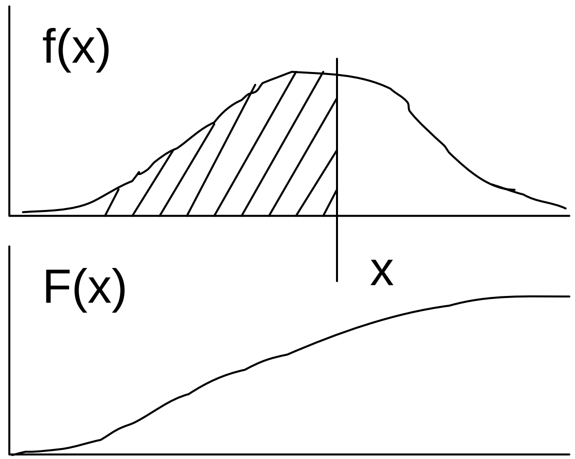
Variables aleatorias continuas

•Una variable aleatoria es continua si toma valores en un intervalo $(x_a,x_b) \subseteq \Re$
Se dice absolutamente continua si $P(x \leq \xi \leq x+dx) = f(x)dx$, donde f es su función de densidad (generaliza el histograma con infinitas clases)

•Propiedades:

1. $f(x) \geq 0, \forall x$
2. $\int_{-\infty}^{+\infty} f(x)dx = 1$
3. $F(x) = P(-\infty \leq \xi \leq x) = \int_{-\infty}^{x} f(t)dt$
4. $f(x) = dF(x)/dx \quad \Rightarrow$

el modelo de probabilidad se define con f o F.



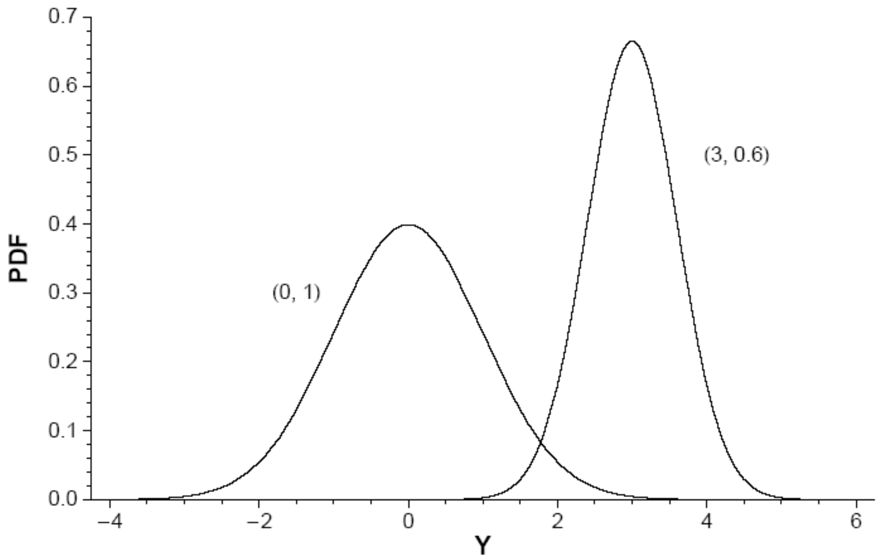
•Los puntos o valores discretos de la variable aleatoria continua no tienen masa de probabilidad. La probabilidad de un valor x=a es la del intervalo [a-1/2,a+1/2]

•La probabilidad de un intervalo [a,b] es $P(a \leq \xi \leq b) = \int_{a}^{b} f(t)dt$

•Variable aleatoria mixta: $F(x) = \alpha F_1(x)+(1 - \alpha)F_2(x)$, $0 \leq \alpha \leq 1$, F_1 vad, F_2 vac

Modelización de la incertidumbre
Variables aleatorias continuas

Normal(A,B) $B > 0$

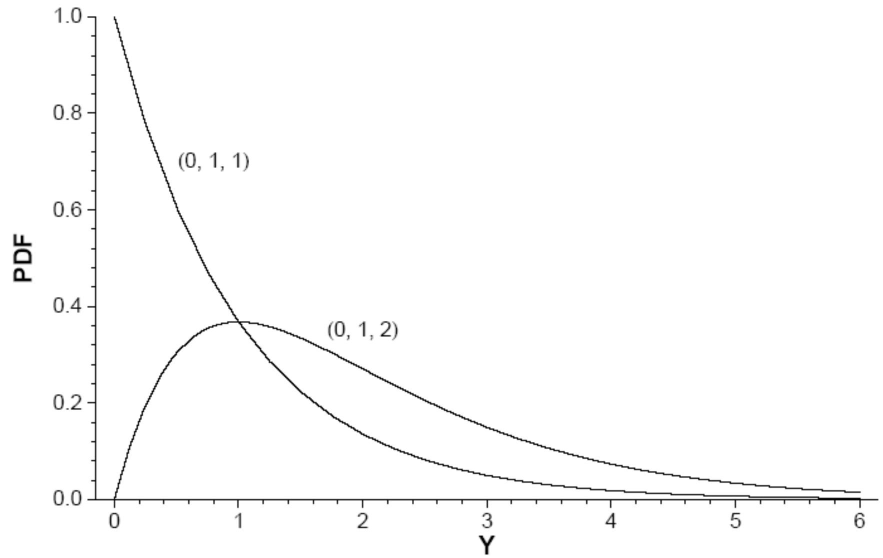


$$\text{PDF} = \frac{1}{B\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{y-A}{B}\right)^2\right)$$

$$\text{CDF} = \Phi\left(\frac{y-A}{B}\right)$$

Parameters -- A (μ): Location, B (σ): Scale

Gamma(A,B,C) $y > A, \quad B > 0, \quad 0 < C \leq 100$



$$\text{PDF} = \frac{1}{B\,\Gamma(C)}\left(\frac{y-A}{B}\right)^{C-1} \exp\left(\frac{A-y}{B}\right)$$

$$\text{CDF} = \Gamma\left(C, \frac{y-A}{B}\right)$$

Parameters -- A (γ): Location, B (β): Scale, C (α): Shape

Modelización de la incertidumbre

Variables aleatorias. Características

•Esperanza E[X], medida de centralización,
promedio de los valores de la variable con su
probabilidad (va discreta) o
densidad de probabilidad (va continua).
Es un operador lineal $E[aX+b]=aE[X]+b$, $E[h(X)] = \int_{-\infty}^{+\infty} h(t)f(t)dt$

•Varianza Var(X),
medida de dispersión asociada a la Esperanza
$Var(X) = \sigma^2 = E[(X-E[X])^2]$,
promedio con su probabilidad (va discreta) o densidad de
probabilidad (va continua). $Var(X) = E[X^2-E[X]^2]$. $Var((aX+b) = a^2Vax(X)$

•Cuantiles de orden $p \in [0,1]$,
valores de la variable aleatoria que son la raiz
de la ecuación $F(x_p) = p$, $p \in \{1/4,1/2,3/4\}$ x_p cuartiles,
$p \in \{1/10,1/5,\ldots 9/10\}$ xp deciles, $p \in \{1/100,1/50,\ldots 99/100\}$ xp percentiles

Modelización de la incertidumbre

Variables aleatorias. Características

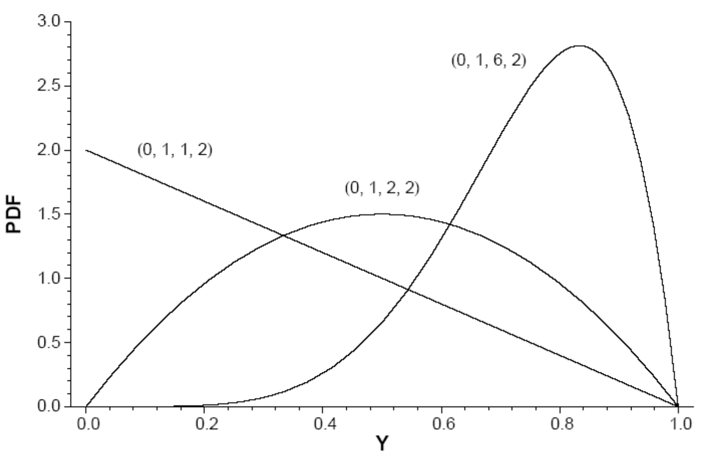
- Momentos de orden k: 0,1,3,4,5,…

$E[X^k]$, $\mu_k = E[(X-E[X])^k]$, $E[(X-x^*)k]$. CAs = μ_3 / σ^3, CAp = $(\mu_4 / \sigma^4) - 3$, CV = σ /

Mediana: cuartil $x_{1/2}$, Moda: Máximo de P_j (va discreta) o de f(x) (va continua

Tipificación: $\forall$ va X, Y=(X-E[X])/Var(X) presenta E[Y]=0 y Var(Y)=1.

Beta(A,B,C,D) $A < y < B, \quad C, D > 0$

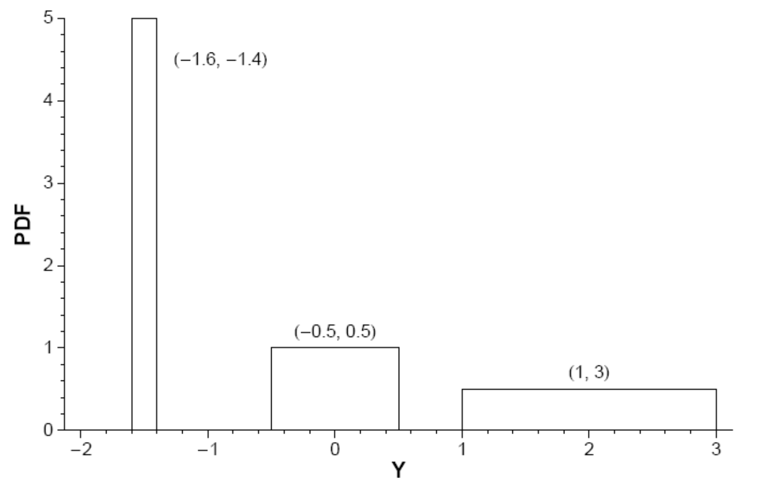


$$\text{PDF} = \frac{\Gamma(C+D)}{\Gamma(C)\,\Gamma(D)\,(B-A)^{C+D-1}}\,(y-A)^{C-1}\,(B-y)^{D-1}$$

$$\text{CDF} = I\left(\frac{y-A}{B-A}, C, D\right)$$

Parameters -- A: Location, B: Scale (upper bound), C, D (p, q): Shape

Uniform(A,B) $A < y < B$



$$\text{PDF} = \frac{1}{B-A}$$

$$\text{CDF} = \frac{y-A}{B-A}$$

Parameters -- A : Location, B : Scale (upper bound)

Modelización de la incertidumbre

Variables aleatorias

- Teorema de Markov

Dada la variable aleatoria ξ y $g(\xi) \geq 0$, $\forall$ $K > 0$,

$P(g(\xi) \geq K) \leq E[g(\xi)]/K$

- Desigualdad de Chebyschev

Dadas $E[\xi]$ y σ de cualquier variable aleatoria, $\forall$ $K > 0$

$P(|\ \xi - E[\xi]| < K\ \sigma) \geq 1 - 1/K^2$

$P(|\ \xi - E[\xi]| \geq K\ \sigma) < 1/K^2$